INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 : C12N 15/00, 1/18, 1/20 C12N 5/00, C12P 21/00	A1	(11) International Publication Number: WO 87/ 03300
		(43) International Publication Date: 4 June 1987 (04.06.87)

(21) International Application Number: PCT/EP86/00675

(22) International Filing Date: 24 November 1986 (24.11.86)

(31) Priority Application Number: 8529014

(32) Priority Date: 25 November 1985 (25.11.85)

(33) Priority Country: GB

(71) Applicant: BIOGEN N.V. [NL/NL]; Pietermaai 15, Willemstad, Curaçoa (AN).

(72) Inventors: ERNST, Joachim, F. ; 156, chemin des Marais, CH-1255 Veyrier (CH). SCHMEISSNER, Ursula ; 20, rue du Conseil-Général, CH-1205 Genève (CH).

(74) Agent: BANNERMAN, David, Gardner; Withers & Rogers, 4 Dyer's Buildings, Holborn, London EC1N 2JT (GB).

(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: ENHANCED SECRETION OF HETEROLOGOUS PROTEINS BY HOSTS USING SUBSTITUTED PROMOTERS

(57) Abstract

Improved secretion of heterologous proteins by hosts such as yeast by using promoters of at most intermediate strength with heterologous DNA secretion signal sequences. A promoter of at most intermediate strength, such as the actin (*ACT*) or iso-l-cytochrome *c* (*CYCl*) promoter in *S. cerevisiae* is operatively linked to a DNA signal sequence, such as the *MFα*1) signal sequence. A DNA sequence for a selected protein, such as somatomedin C (SMC), tissue plasminogen activator (TPA) or tumor necrosis factor (TNF), may be operatively linked to the DNA signal sequence.

ENHANCED SECRETION OF HETEROLOGOUS PROTEINS BY HOSTS USING SUBSTITUTED PROMOTERS

Technical Field Of The Invention

5   This invention relates to expression systems and recombinant DNA molecules that facilitate enhanced secretion of heterologous proteins by hosts, to hosts comprising such recombinant DNA molecules and to methods of producing desired proteins using such
10   hosts.

Background Of The Invention

       Proteins prepared by recombinant DNA methods are sometimes difficult to isolate because the protein must be extracted from the transformed microorganism
15   by means such as cell lysis that typically destroy the microorganism. Obtaining sufficiently pure protein in high yield from such extraction procedures is difficult because of the large number of different organic compounds liberated when lysis
20   takes place.

       One attractive alternative to typical separation procedures is to have the host secrete the selected protein into its environment. A secreted protein is more easily separated from the culture
25   medium, and the microorganism can survive the separation process to produce and to secrete more of the desired protein.

-2-

Most proteins naturally secreted by pro-
karyotes and eukaryotes are initially synthesized in
the form of precursors containing an amino-terminal
extension several amino acids long. This extension,
called a secretion leader or signal sequence, allows
the precursor protein to cross the cell membrane of
the microorganism and enter the culture medium or
periplasmic space of the cell. During secretion,
the secretion leader is cleaved from the protein,
leading to the presence of mature protein in the
culture medium or periplasmic space.

Although the rate of secretion in bacteria
is typically very low, bacterial secretion has been
used with recombinant DNA technology. See, for
example, United States Patents 4,411,994 and 4,338,397,
and Villa-Komaroff et al., Proc. Natl. Acad. Sci. USA,
75, 3727-31 (1978).

The general safety of yeasts and human
experience with yeast fermentation have made yeasts
desirable candidates for use as hosts in recombinant
DNA technology. However, reported attempts to obtain
secreted mature proteins from recombinant yeasts
have suffered from low yields. See, for example,
C. N. Chang et al., "Recognition and Cleavage of
Hybrid Invertase Signals and Mature Forms of Human
Interferon (IFN-α2) in Yeast," Meeting Abstracts,
The Molecular Biology of Yeast, Cold Spring Harbor
Laboratory, Cold Spring Harbor, New York, p. 393
(1983); B. Meyhack and A. Hinnen, "High Levels of
Expression of Foreign Genes Under the Control of the
Yeast PHO5 Promoter," Meeting Abstracts, The Molecular
Biology of Yeast, Cold Spring Harbor Laboratory,
Cold Spring Harbor, New York, p. 156 (1983); S. D.
Emr, "An MFα1-SUC2 (α-Factor-Invertase) Gene Fusion
for Study of Protein Localization and Gene Expression
in Yeast," Proc. Natl. Acad. Sci. USA, 80, pp. 7080-84
(1983); A. Brake et al., "α-Factor-Directed Synthesis

-3-

and Secretion of Mature Foreign Proteins in
Saccharomyces cerevisiae," Proc. Natl. Acad. Sci.
USA, 81, pp. 4642-46 (1984); G. A. Bitter et al.,
"Secretion of Foreign Proteins from Saccharomyces
5   cerevisiae Directed by α-Factor Gene Fusions," Proc.
Natl. Acad. Sci. USA, 81, pp. 5330-34 (1984);
A. Singh, et al., "Synthesis, Secretion and Processing
of α-Factor-Interferon Fusion Proteins in Yeast,"
Nucl. Ac. Res., 12, pp. 8927-38 (1984); European
10   Patent Application 123,228; and European Patent
Application 128,733; and G. Bitter et al, "Secretion
of Foreign Proteins from Saccharomyces cerevisiae
Directed by α-Factor Pheromone," Nucl. Ac. Res., 11,
pp. 4049-63 (1983).

15   Summary of the Invention

        The present invention solves the foregoing
problems by providing expression systems that facili-
tate enhanced secretion of heterologous proteins.
We have found that using a promoter of at most inter-
20   mediate strength and a heterologous DNA secretion
signal sequence in a multicopy vector results in a
higher yield of a desired protein from the host com-
pared to hosts transformed by a multicopy vector
using a strong promoter.
25   Thus, one embodiment of the present inven-
tion comprises a DNA sequence that is an expression
and secretion control sequence for producing and
secreting a selected protein from a host in which
the protein is made, said sequence comprising (a) a
30   promoter that is active in said host and is of at
most intermediate strength, and (b) a heterologous
DNA secretion signal sequence, recognized by said
host, beginning with a start codon, operatively linked
to said promoter. The expression and secretion con-
35   trol sequence may then be operatively linked to a
DNA sequence coding for a desired protein and employed

-4-

to transform a host to produce and secrete the desired
protein. Yeast is a preferred host, and the actin
(ACT) and iso-1-cytochrome c (CYC1) promoters are
preferred promoters for yeast hosts. S. cerevisiae
5   is the most preferred host. When the host is
S. cerevisiae, the secretion signal sequence is
preferably the secretion signal sequence of the
MFα1 gene.

The present invention also relates to hosts
10   transformed by the recombinant DNA molecules discussed
above and to methods of using transformed hosts in the
preparation and secretion of desired proteins.

Brief Description Of The Drawings

Figure 1A shows the structure of the MFα1
15   secretion precursor. Stippled regions indicate spacer
regions separating the MFα1 repeats. During secretion
the secretion leader is cleaved off at the position
indicated by the arrow.

Figure 1B shows a general scheme for the
20   construction of a yeast secretion precursor comprising
the alpha mating factor (MFα1) secretion leader fused
to the desired heterologous protein. During secre-
tion, the secretion leader is cleaved at the position
indicated by the arrow.

25   Figure 2 shows the construction of a gene
fusion between the MFα1 expression control sequence
and the SMC gene and insertion of the MFα1/SMC fusion
into a yeast expression vector.

Figure 3 shows two graphs: growth (as
30   reflected by optical density at 600 nm) and SMC
levels in the culture fluid of yeast strain BJ1991
transformed with URA3-type expression vectors. The
CYC1 promoter construction (p336/1) is shown by (●);
the ACT promoter construction (p364/1) by (▲); and
35   the MFα1 promoter construction (p446/1) by (■).

-5-

Figure 4A shows the construction of actin promoter fragments ending in EcoRI sites.

Figure 4B shows the DNA structure of promoter ends for actin and the pEX-5, pEX-7, and pEX-8 vectors.

Figure 5 shows the construction of SMC secretion vectors based on the ACT and CYC1 promoters.

Figure 6 shows the structure of SMC secretion vectors based on the yeast LEU2 gene.

Figure 7 shows two graphs: growth (as reflected by optical density at 600 nm) and SMC levels in the culture fluid of yeast strain BJ1991 transformed with LEU2-type expression vectors. The CYC1 promoter construction (504/1) is shown by (●); the ACT promoter construction (p482/18) by (▲); and the MFα1 promoter construction (pMF-SMC, rearranged) by (■).

Figure 8 shows the construction of secretion vectors for TNF.

Figure 9 shows the construction of secretion vectors for TPA.

Detailed Description Of The Invention

In order that the invention may be more fully understood, the following detailed description is provided. In this specification, some of the following terms are used:

Cloning -- the process of obtaining a population of organisms or DNA sequences derived from one such organism or sequence by asexual reproduction.

Recombinant DNA Molecule or Hybrid DNA -- A molecule consisting of sequences of DNA from different genomes which have been joined end-to-end outside of living cells and that can be maintained in living cells.

Protein -- A polypeptide containing a linear series of more than fifty amino acids, e.g., pro-

insulin, serum albumin, human growth hormone, para-
thyroid hormone, and interferon. As used herein,
however, a protein also comprises a polypeptide chain
of fewer than fifty amino acids.

5   Polypeptide -- A linear series of amino
acids connected one to the other by peptide bonds
between the amino and carboxy groups of adjacent
amino acids.

           Expression -- The process undergone by a
10   gene to produce a polypeptide or protein. It is a
combination of transcription and translation.

           Transcription -- The process of producing
mRNA from a gene.

           Translation -- The process of producing a
15   protein or polypeptide from mRNA.

           Promoter -- The region of DNA responsible
for binding RNA polymerase to initiate transcription.
In bacterial expression systems a promoter is located
before the ribosome binding site. A strong promoter
20   is defined as one that has a strong affinity for RNA
polymerase and that would accordingly be expected to
aid in obtaining high expression rates. A promoter
of intermediate strength has a lower affinity for
RNA polymerase.

25   In recombinant systems, the strength of a
promoter may be affected by the host, the desired
protein and the expression vector of the recombinant
system. The effects of these factors are more fully
discussed below. In systems that do not secrete the
30   desired protein, i.e., intracellular expression
systems, levels of mRNA and levels of expressed pro-
tein are proportional to promoter strength. In
intracellular expression systems, no changes of
plasmid copy number have been observed with variation
35   of promoter strength. J. Mellor et al, "Factors
Affecting Heterologous Gene Expression in
Saccharomyces cerevisiae," Gene, 33, pp. 215-26

-7-

(1985). In secretion expression systems, we have
discovered that promoter strength may affect the
number of copies of a multicopy expression vector
produced within a given host. In secretion expres-
5   sion systems, a host will produce many copies of a
multicopy vector that contains an intermediate or
weak promoter but will produce fewer copies of a
multicopy vector that contains a strong promoter.
An estimate of the strength of a promoter may be
10   based on the percent of total mRNA in a cell
produced by the expression control sequences of that
promoter. Table 1 shows a list of genes, including
their associated promoters and associated mRNA, as
described in the literature. A characterization of
15   the promoter strength is also included.

-8-

TABLE 1

Gene	Protein	Percent of Total mRNA	Reference	Characterization of Promoter Strength
pgap 63	G3PDH	1.5-6	1	very strong
peno 8	enolase	1-3	1	very strong
ADH-I	alcoholde-hydrogenase	0.7-2	1	very strong
	histone 2B	0.4	1	strong
tcm	ribosomal protein P1	0.3	2, 3	strong
act	actin	0.15	3	intermediate
cyc1	iso-1-cytochrome c	0.05	1	weak
URA3	-	0.008	2	very weak
cyc7	iso-2-cytochrome c	0.003	1	very weak
trp1	-	<0.01	4	very weak

_____

1.   J. L. Bennetzen and B. D. Hall, "Codon Selection
     in Yeast," J. Biol. Chem., 257, pp. 3026-31 (1982).

2.   C. H. Kim and J. R. Warner, "Messenger RNA for
     Ribosomal Proteins in Yeast," J. Mol. Biol.,
     165, pp. 78-89 (1983).

3.   H. J. Himmelfarb et al., "Isolation of the SUP45
     Omnipotent Suppressor Gene of Saccharomyces
     cerevisiae and Characterization of its Gene
     Product," Mol. Cell. Biol., 5, pp. 816-22 (1985).

4.   M. J. Dobson et al., "Expression in Saccharomyces
     cerevisiae of Human Interferon-Alpha Directed
     by TRP1 5' Region," Nucl. Ac. Res., 11,
     pp. 2287-2302 (1983).

Generally, promoters of glycolytic genes
are considered "strong." Those skilled in the art
will understand, however, that promoters may not
have the same strength within different strains of a
5   host, with different desired proteins or in different
multicopy vectors. In S. cerevisiae, for example,
the MFαl promoter is active in MATα strains, but
inactive in MATa strains. Those skilled in the art
will appreciate that promoter strength depends upon
10   many factors. Merely because a promoter is unaccept-
able with respect to one combination of host, vector
and desired protein does not mean that the promoter
will not be acceptable for another combination of
host, vector and desired protein.

15   Promoters of the present invention should
be selected so that the combination of promoter
strength and copy number of the multicopy vector pro-
vides optimum yields of desired protein. Promoters
of intermediate strength provide the best combination
20   of promoter strength with copy number. Weaker pro-
moters, in combination with high copy numbers, are
also within the scope of the invention. Strong
promoters adversely affect the copy number of the
multicopy vectors within the transformed host, and
25   unacceptable yields are obtained from those trans-
formants. A promoter should be selected that has
maximal strength without significantly reducing
plasmid copy number. Generally, a promoter may be
considered of at most intermediate strength with
30   respect to a given combination of host, vector and
desired protein if the percent of total mRNA in the
host corresponding to the promoter is less than 0.3%,
and preferably is at most 0.15%.

         Ribosome Binding Site -- The region of DNA
35   which codes for a site on mRNA which helps the mRNA
bind to the ribosome, so that translation can begin.
In bacterial expression systems, a ribosome binding

site is located after (downstream from) the promoter
and before (upstream from) the translational start
signal of the DNA sequence to be expressed to produce
the desired protein.

5   Gene -- A DNA sequence which encodes, as a
template for mRNA, a sequence of amino acids charac-
teristic of a specific polypeptide or protein.
           Expression Control Sequence -- A DNA
sequence that controls and regulates expression of
10   genes when operatively linked to those genes. Such
sequences include the lac system, the β-lactamase
system, the trp system, the tac, and trc systems,
the major operator and promoter regions of phage λ,
the control region of fd coat protein, the early and
15   late promoters of SV40, promoters derived from polyoma
virus and adenovirus, metallothionine promoters, the
promoter for 3-phosphoglycerate kinase or other glyco-
lytic enzymes, the promoters of acid phosphatase,
e.g., Pho5, the promoters of the yeast α-mating
20   factors, and other sequences known in the art to
control the expression of genes in prokaryotic or
eukaryotic cells and their viruses or combinations
thereof. For mammalian cells, the gene can be linked
to a eukaryotic promoter, such as that for the SV40
25   early region, coupled to the gene encoding dihydro-
folate reductase and selectively amplified in Chinese
hamster ovary cells to produce a cell line contain-
ing many copies of actively transcribed eukaryotic
genes. Those skilled in the art will appreciate
30   that not every expression control sequence listed
above as an example may be suitable for use with
each host, desired protein and vectors combination.
           Percursor of a Protein -- A protein with
a signal sequence operatively linked to the protein.
35   Typically, the precursor is synthesized within a host
cell, e.g., preproinsulin, preserum albumin, prehuman
growth hormone, preparathyroid hormone, and preinter-

-11-

feron. In accordance with this invention, a mature
protein is obtained by secreting the precursor through
the cell membrane of a host with an attendant loss
or clipping of the signal sequence of its precursor.

5   Nucleotide -- A monomeric unit of DNA or
RNA consisting of a sugar moiety (pentose), a phos-
phate, and a nitrogenous heterocyclic base. The
base is linked to the sugar moiety via the glycosidic
carbon (1' carbon of the pentose) and that combina-
10   tion of base and sugar is called a nucleoside. The
base characterizes the nucleotide. The four DNA
bases are adenine ("A"), guanine ("G"), cytosine
("C") and thymine ("T"). The four RNA bases are A,
G, C and uracil ("U").

15   DNA Sequence -- A linear series of nucleo-
tides connected one to the other by phosphodiester
bonds between the 3' and 5' carbons of adjacent
pentoses.

            Codon -- A DNA sequence of three nucleo-
20   tides (a triplet) which encodes, through messenger
RNA ("mRNA"), an amino acid, a translational start
signal or a translational termination signal. For
example, the nucleotide triplets TTA, TTG, CTT, CTC,
CTA and CTG encode for the amino acid leucine ("Leu"),
25   TAG, TAA and TGA are translational stop signals and
ATG is a translational start signal.

            Plasmid -- A non-chromosomal double-stranded
DNA sequence comprising an intact "replicon" such
that the plasmid is replicated in a host cell. When
30   the plasmid is placed within a host organism, the
characteristics of that organism are changed or trans-
formed as a result of the DNA of the plasmid. For
example, a plasmid carrying the gene for tetracycline
resistance (Tet^R) transforms a host cell previously
35   sensitive to tetracycline into one which is resistant
to it. A host cell transformed by a plasmid is called
a "transformed host" or a "transformant."

-12-

Phage or Bacteriophage -- Bacterial virus
which may include DNA sequences contained in a protein
envelope or coat ("capsid").

Cloning Vehicle -- A plasmid, phage DNA or
other DNA sequence which is able to replicate in a
host cell, characterized by one or a small number of
endonuclease recognition sites at which its DNA
sequence may be cut in a determinable fashion without
attendant loss of an essential biological function
of the DNA, e.g., replication, production of coat
proteins or loss of promoter or binding sites, and
which contains a marker suitable for use in the iden-
tification of transformed cells, e.g., tetracycline
resistance or ampicillin resistance. A cloning
vehicle is also known as a vector. In the present
invention, the vector must be a "multicopy" vector,
i.e., capable of producing multiple copies of itself
within the host. Single copy vectors, incapable of
producing copies, are not within the scope of this
invention.

Host -- An organism which on transformation
by a cloning vehicle enables the cloning vehicle to
replicate and to accomplish its other biological
functions, e.g., the production of polypeptides or
proteins through expression of the genes of a plasmid.

DNA Signal Sequence -- A DNA sequence which
encodes, as a template for mRNA, a sequence of
typically hydrophobic amino acids at the amino
terminus of the polypeptide or protein, i.e., a
"signal sequence" or "secretion leader sequence" of
the protein. For secretion, such a DNA signal
sequence is operatively linked to and located immedi-
ately before the DNA sequence of the desired protein
and after its translational start signal (e.g., ATG).
It is believed that only a portion of a signal
sequence of a precursor of a protein is essential
for the precursor of the protein to be transported

through the cell membrane of a host and for proper
clipping of the precursor's signal sequence to free
the protein during secretion. Combinations of signal
sequences or parts of signal sequences may also be
used provided that proper processing takes place in
the host. Hence, the term "DNA signal sequence" as
used herein includes those DNA sequences that code
for that portion of a signal sequence required for
secretion.

In accordance with this invention, to obtain
expression of a desired protein, an expression system
of this invention is operatively linked to a DNA
sequence coding for the desired protein and used to
transform an appropriate host. The host may then be
cultured under appropriate conditions of growth.
The desired protein may then be isolated from the
culture. Optimum results will, of course, depend on
a number of variables, including the appropriate
choice of host, promoter, vector, and growth condi-
tions.

Many hosts have been used in recombinant
DNA technology and are well known in the art. A
wide variety of hosts may be useful in the method of
this invention. These hosts include for example,
bacteria, such as E.coli (for example E.coli HB101
or E.coli MC1061), Bacillus, Streptomyces, and
Pseudomonas, fungi, such as yeasts, animal cells,
such as CHO cells, mouse, swine, bovine, fowl or
fish cells, plant cells in tissue culture, human
tissue cells, or other hosts known in the art.

The selection of an appropriate host is
controlled by a number of factors recognized by the
art. These include, for example, compatibility with
the chosen vector, ease of recovery of the desired
protein, expression characteristics, biosafety and
costs. In the present invention, the host must be
able to recognize both the promoter used and

-14-

the secretion signal sequence used.   In addition,
the toxicity of the desired protein on the host must
be considered, as well as the viability of the host
when transformed by a muticopy vector.   No absolute

5   choice of host may be made for a particular recom-
binant DNA molecule or polypeptide from any of these
factors alone.   Instead, a balance of these factors
must be struck with the realization that not all
hosts may be equally effective for expression of a

10   particular DNA sequence operatively linked to a
particular expression control sequence.   While the
present invention is not limited to _S. cerevisiae_,
yeast is a preferred host, and _S. cerevisiae_ is
especially preferred.

15   Various promoters may be used in the
expression systems of this invention, provided, how-
ever, that the promoter is active in the chosen host
and is of at most intermediate strength.
          The toxicity of the desired protein to the

20   host affects the promoter choice; for secretion of a
less toxic protein, a relatively stronger promoter
can be used than for secretion of a more toxic pro-
tein.   The coding region attached to a promoter in a
multicopy vector also affects the strength of the

25   promoter.   For example, the pgk promoter is very
strong when attached to its own coding region (PGK),
but weak when attached to an alpha interferon gene.
J. Mellor et al., "Factors Affecting Heterologous
Gene Expression in _Saccharomyces cerevisiae_," _Gene_, _33_,

30   pp. 215-26 (1985).   In choosing a promoter it should
also be appreciated that the junction of a strong
promoter to a heterologous gene will considerably
weaken the strong promoter, as described by Mellor
et al., _supra_.   Because promoter strength is affected

35   by so many variables, the fact that a promoter is
too strong for use in one host-vector-desired protein

-15-

combination does not exclude the use of that promoter
in a different combination.

Promoters that may be considered for use
in recombinant DNA molecules made from the present
5   invention include, for the yeast S. cerevisiae, the
promoters of the ACT (actin) gene, the CYC1 (iso-1-
cytochrome c) gene and the URA3 gene. Kim and Warner,
J. Mol. Biol., 165, pp. 79-89 (1983) (mRNA level
about 0.008%). For secretion expression systems,
10   strong promoters not suitable for use with
S. cerevisiae include, for example, the G3PDH and
the enolase promoters.

Preferred promoters in secretion expression
systems for use with bacteria such as E.coli and
15   B. subtilis should be of at most intermediate strength.
Strong promoters, such as the trp, trc, lac promoters
should generally be avoided. Intermediate promoter
strength might, however, conveniently be adjusted
with regulated promoters, such as the trp or lac
20   promoters, if intermediate inducing growth conditions
are chosen.

In the examples, we have used promoters
that are active during all phases of growth. If the
promoters were to be shut off during growth, the
25   adjustment of vector copy number would presumably
not take place unless the promoters were induced.
Regulated promoters, however, would probably not
provide an advantage. Presumably, turning on a strong
promoter in a high-copy secretion construction would
30   immediately lyse the host, as in insulin secretion
by E.coli.

As discussed above, we have discovered that,
for secretion expression systems, a host makes few
copies (i.e., low copy numbers) of multicopy vectors
35   if the promoter is strong in the host-vector-desired
protein system, and more copies (i.e., higher copy
numbers) are made of multicopy vector having a weaker

-16-

promoter, resulting in improved yields of desired
protein. However, the use of a promoter weaker than
the CYC1 promoter discussed in Example 3 (e.g., CYC7
or trp1) will not improve yields due to higher copy
5   numbers, because maximal copy numbers are seen for
the CYC1 promoter construction (Table 2). Lowest
strength promoters are limited by the promoter
strength, and the strongest promoters give only
limited copy numbers. Though not wishing to be bound
10   by any theory, it appears that high copy numbers
combined with strong promoters result in high levels
of mRNA, oversecretion of the desired protein or
both. In a population of transformants containing
cells transformed by vectors having strong promoters,
15   a mixture of copy numbers occurs in the population.
Those cells having high copy numbers with strong
promoters will stop growing and eventually lyse,
while those transformants with lower copy numbers
and the same strong promoter will survive. The
20   population will eventually contain only low copy
number transformation. See generally T. J. Silhavy,
et al., "Mechanisms of Protein Localization,"
Microbiol. Res., 47, pp. 313-34 (1983).
            The multicopy vector and, in particular,
25   the sites chosen therein for insertion of the expres-
sion systems of this invention are determined by a
variety of factors, e.g., number of sites susceptible
to a particular restriction enzyme, size of the pro-
tein to be expressed, expression characteristics
30   such as the location of start and stop codons relative
to the vector sequences, and other factors recognized
by those of skill in the art. The choice of a vector
is determined by a balance of these factors, not all
selections being equally effective for a given case.
35   The preferred vectors are plasmids, such as those
containing origins of replication derived from
chromosomal and non-chromosomal DNA. In the yeast

-17-

S. cerevisae, the various ARS- and 2 μ-type vectors
(Botstein et al.) are preferred. In E.coli, multi-
copy plasmids, such as ColE1-, pBR322 and RP4 plasmids,
or M13 and lambda phage vectors, can be used. For
5   animal cells, non-integrating multicopy vectors,
such as vectors based on the bovine papilloma virus,
are preferred. The choice of vector will also be
influenced by the chosen promoter. optimal promoter
and vector systems will lead to less toxicity of the
10   secreted protein (due to too high copy number or
over production based on too strong a promoter) and
consequently, higher levels of secretion may be
obtained.

         The heterologous secretion signal sequences
15   of this invention must be recognized and correctly
processed by the host. The signal sequence may com-
prise a naturally occurring signal sequence, an effec-
tive portion of a naturally occurring signal sequence,
or a combination of two or more signal sequences or
20   effective portions of signal sequences. The signal
sequence must be operatively linked to both the
promoter and the start signal, ATG. Preferably, the
signal sequence begins with the translation start
signal. In a preferred embodiment, the signal
25   sequences is selected from signal sequences already
used by the host to aid secretion of its own
proteins.

         The desired protein may include any poly-
peptide or protein, and may include fusions, prepro-
30   teins, immature proteins or any desired sequence of
amino acids. Preferably, however, desired proteins
of this invention include proteins, and other amino
acid sequences, that are secretable by some micro-
organism or cell. Examples of such secretable
35   proteins and other amino acid sequences include
serum proteins, analgesic polypeptides such
β-endorphin, somatostatin, insulin, growth hormone

-18-

(human and bovine), luteinizing hormone, ACTH,
pancreatic polypeptide preproteins, preproinsulin,
proinsulin, and the A and B chains of insulin. The
desired proteins of the present invention may comprise
5   proteins and other amino acid sequences naturally
secreted into the medium by the selected host.

    The desired protein of the present inven-
tion need not be secreted by the selected host. As
will be seen in Example 6, the host need only manu-
10 facture the desired protein operatively linked to
the chosen signal sequence and correctly process
the desired protein into the endoplasmic reticulum.
If the desired protein enters the secretion pathway
of the host, benefits may be obtained for the desired
15 protein, such as glycosylation. Secretion is, how-
ever, the preferred result for the desired protein.

    The present invention may be useful in
mammalian cells, if non-integrating, multicopy
vectors, such as vectors derived from bovine
20 papilloma virus (D. DiMaio et al., "Bovine Papilloma-
Virus Vector that Propagates as a Plasmid in Both
Mouse and Bacterial Cells," Proc. Nat. Acad. Sci.
USA, 79, pp. 30-34 (1982)) are used in conjunction
with an intermediate strength promoter.

25   The following, non-limiting examples will
serve to further illustrate the present invention.

Example 1

Preparation Of MFα1/SMC DNA Sequences

    This example describes the preparation of
30 a transformed yeast which is used in a later example
for comparative purposes. The transformed yeast of
this example has a plasmid containing a MFα1 promoter
(which is a strong promoter in yeast as shown below)
operatively linked to a MFα1 secretion leader and a
35 DNA sequence coding for SMC. Other examples describe

-19-

the substitution of the ACT and CYC1 promoters for
the MFα1 promoter, and a comparison of the secretion
levels of the variously transformed yeasts. In addi-
tion, other examples describe the preparation of
5 transformed yeasts having plasmids characterized by
DNA sequences coding for desired proteins other than
SMC.

We first isolated the gene for pre-MFα1
and its associated expression control systems and we
10 then operatively linked the MFα1 promoter and signal
sequence to a DNA sequence for SMC.

The DNA sequence coding for the precursor
of MFα1 has been sequenced by J. Kurjan and
I. Herskowitz, "Structure of a Yeast Pheromone Gene
15 (MFα): A Putative α-Factor Precursor Contains Four
Tandem Copies of Mature α-Factor," Cell, 30,
pp. 933-43 (1982). We isolated the gene coding for
those pre-MFα1 and its associated expression control
sequences from S. cerevisiae using the library con-
20 structed by K. A. Nasmyth, and S. I. Reed, "Isolation
of Genes by Complementation in Yeast: Molecular
Cloning of a Cell-Cycle Gene," Proc. Natl. Acad. Sci.
USA, 77, 2119-23 (1980). Instead of using the cdc28
mutant described in Nasmyth and Reed for isolation
25 and cloning of the cdc28 gene, we used the following
oligonucleotide corresponding to amino acids 97 to
102 of the published pre-MFα1:

(5') GTA CAT TGG TTG C/GCC G/A/TGG (3')

The MFα1 secretion precursor, comprising a
30 secretion leader and four MFα1 repeats is shown in
Figure 1A. The stippled regions indicate spacer
regions between the MFα1 repeats. Before substituting
a desired heterologous protein for the four MFα1
repeats, as shown in Figure 1B, the pre-MFα1 gene
35 and associated expression control sequence, which
resided on a 1.7 kb EcoRI fragment, were subcloned
into pUC18, the selected cloning vector. The result-

-20-

ing plasmid (p220/3) now included an a MFα1 promoter, an MFα1 signal sequence and an MFα1 DNA coding sequence. In order to isolate the MFα1 coding sequence from the MFα1 promoter and MFα1 signal

5   sequence, we subjected plasmid p220/3 to mutagenesis using the procedure described in B. A. Oostra et al., "Transforming Activity of Polyoma Virus Middle-T Antigen Probed by Site-Directed Mutagenesis," Nature, 304, pp. 456-59 (1983). As a result of the muta-

10 genesis, a convenient BglII site was introduced at the junction between the secretion leader and the coding sequence for the repeats in the MFα1 DNA sequence as follows:

modified
15 oligomer: (3') TA TTT TCT CTA GAA CTT CGA ACC (5')

    original: (3') TA TTT TCT CTC CGA CTT CGA ACC (5')
    sequence   lys arg glu ala glu ala trp

        The resulting plasmid was called p254
    (Figure 2). It was further modified to prevent
20 unintended cleavage at a BglII site found at the 5'
    end of the MFα1 promoter. We altered this undesired
    BglII site by cleaving with BglII and filling in the
    protruding ends using the large (Klenow) fragment of
    DNA polymerase I and deoxynucleotide triphosphates
25 and subsequent ligation using T4 ligase. To allow
    for a different cleavage in the plasmid (Examples 5
    and 6), a StuI site was introduced, in a procedure
    similar to the mutagenesis described above, at the
    position corresponding to the same lys-arg cleavage
30 site:

    modified
    oligomer: (5') TA AAA AGG CCT CTT GAA GC (3')

    BglII-sequence: (5') TA AAA AGA GAT CTT GAA GC (3')
                           lys arg

35   We then made a MFα1/SMC (secretion leader/
    desired protein) fusion by inserting a 500 bp HindIII
    fragment carrying a synthetic SMC gene starting with

-21-

a unique NcoI site into the HindIII site of plasmid
p254 (described above) which contained the MFαl pro-
moter and secretion leader and obtained plasmid F-9
as shown in Figure 2. The synthetic SMC gene is
5   described in G. Buell et al., "Optimizing the Expres-
sion in E.coli of a Synthetic Gene Encoding
Somatomedin-C (IGF-1)," Nucl. Ac. Res., 13, pp. 1923-
38 (1985). We then cut plasmid F-9 with BglII and
NcoI, with simultaneous S1 treatment and religation
10   to remove the Δ portion of plasmid F-9 and obtain
pS30/25, shown in Figure 2. The correct fusion of
the MFαl secretion leader and the SMC structural DNA
sequence had the following sequence (glycine is the
first amino acid of SMC):
15   (secretion leader)-AAA-AGA-GGT-CCA-(SMC)
                              lys arg gly pro
            As a result the SMC DNA sequence was opera-
tively linked to the MFαl secretion leader so that
the SMC protein would be properly secreted by yeast.
20   As outlined in Figure 2, we introduced the
MFαl/SMC fusion into a yeast shuttle vector carrying
origins of replication for E.coli and yeast (ori and
the 2 μ origin of replication, respectively), as well
as selectable markers for both organisms (E.coli:
25   bla; yeast: URA3). SMC secretion rates by
S. cerevisiae transformed with this expression vector
are shown in Figure 3.

                        Example 2

          Preparation Of ACT/MFαl/SMC DNA sequences

30   The DNA sequence encoding actin in yeast
is described in D. Gallwitz and I. Sures, "Structure
of a Split Yeast Gene: Complete Nucleotide Sequence
of the Actin Gene in Saccharomyces cerevisiae," Proc.
Natl. Acad. Sci. USA, 77, pp. 2546-50 (1980). Levels
35   of the mRNA coding for actin (D. Gallwitz et al.,

-22-

"The Actin Gene in the Yeast Saccharomyces cerevisiae:
5' and 3' End Mapping, Flanking and Putative Regula-
tory Sequences," Nucl. Ac. Res., 9, pp. 6339-50
(1981)) suggest that the ACT promoter is a promoter
5   of intermediate strength, see Table 1. See also
Himmelfarb, supra.

         In order to use the ACT promoter in the
constructions of this and the following examples, we
10   placed convenient restriction sites at the end of the
promoter. Plasmid pYA301, shown in Figure 4, contains
the ACT gene and its associated expression control
sequences on a 4 kb EcoRI-BamHI fragment. Plasmid
pYA301 is a 4 kb EcoRI-BamHI subclone of pYA208
15   (Gallwitz and Sures, supra) inserted into pBR322.
We cut plasmid pYA301 at the single XhoI site and
treated it with Bal31. Following an empirically
determined time of Bal31 treatment, we cut the plas-
mid with EcoRI and its protruding ends were filled
20   in using the large (Klenow) fragment of DNA poly-
merase I. Finally, the treated plasmid was reli-
gated under dilute conditions. Approximately half
of the resulting plasmids, identified collectively
as pΔ4, contained an EcoRI site. These EcoRI sites
25   had been formed by joining the filled-in EcoRI end
to the Bal31 ACT promoter end. Structures of various
pΔ4 promoter ends in the vicinity of the ATG start
codon of the religated plasmid are shown in Figure 4B.
         To operatively link the isolated ACT
30   promoter to the MFα1 signal sequence SMC fusions, we
matched the 3' end of the ACT promoter to the amino
terminal end of the MFα1/SMC fusion. We accomplished
this by introducing an EcoRI site upstream (i.e.,
earlier in the transcription reading direction) of
35   the ATG start codon of the MFα1 secretion leader by
carrying out the following mutagenesis using the
procedure described in Oostra et al., supra:

modified
oligomer: (5') A ATA TAA ACG AAT TCA AGA ATG AG (3')

original
sequence: (5') A ATA TAA ACG ACC AAA AGA ATG AG (3')

5   We named the resulting plasmid p269/20.

        We joined the BamHI-EcoRI fragment of pEX-7
(Fig. 4) isolated above and containing the ACT pro-
moter to the EcoRI-SalI fragment of plasmid p269/20
(which carried the MFα1 secretion leader) and inserted
10   the fragment into a BamHI and SalI cut pUC8 vector as
shown in Figure 5. We further modified the resulting
plasmid (p309/1) by replacing the MFα1 secretion
leader region of plasmid p309/1 by the secretion
leader region, joined to the SMC gene, of plasmid
15   pS30/25. Plasmid pS30/25 is shown in Figure 2. The
modification was carried out by cutting plasmid
p309/1 with PstI and HindIII. The large fragment was
isolated and joined to a fragment of pS30/25 resulting
from partial digestion of that plasmid with PstI and
20   HindIII. The resulting plasmid, p355/1, (ACT
promoter/MFα1 secretion leader/SMC structural DNA
sequence) was transferred to a yeast shuttle vector,
as shown in Figure 5, which resulted in formation of
plasmid p364/1. SMC secretion rates by yeast trans-
25   formed with this plasmid are shown in Figure 3.

                       EXAMPLE 3

        Preparation Of CYC1/MFα1/SMC Plasmids

        Under inducing conditions, iso-1-cytochrome
c (CYC1) amounts to approximately 0.2% of the total
30   cellular protein of yeast, and the encoding mRNA
amounts to about 0.05% of the total poly(A) RNA.
R. S. Zitomer and B. D. Hall, "Yeast Cytochrome c
Messenger RNA. In Vitro Translation and Specific
Immunoprecipitation of the CYC1 Gene Product,"

-24-

<u>J. Biol. Chem.</u>, <u>251</u>, 6320-26 (1976). The <u>CYCl</u> pro-
moter is thus a weak promoter.

      In order to construct a <u>CYCl</u> promoter/<u>MFαl</u>
secretion leader/SMC coding sequence, we isolated an

5   EcoRI-HindIII fragment carrying the <u>MFαl</u> secretion
leader correctly fused to the SMC gene from plasmid
p355/1, which is shown in Figure 5. We inserted
this fragment between the <u>EcoRI</u> and <u>HindIII</u> sites of
pEX-2 to obtain plasmid p446/1 (See Figure 5).

10   pEX-2 is described in J. F. Ernst and
R. C. Chan, "Characterization of <u>S. cervisiae</u> Mutants
Supersensitive to Aminoglycoside Antibiotics,"
<u>J. Bacteriol.</u>, <u>163</u>, pp. 8-14 (1985).

      The vector p446/1 expressed the <u>MFαl</u>/SMC

15   fusion under the control of the <u>CYCl</u> promoter. SMC
secretion rates from yeast transformed by this vector
are shown in Figure 3.

<center>• <u>EXAMPLE 4</u></center>

<center>Comparison Of Secretion Of SMC Using The
<u>MFαl</u>, <u>ACT</u> And <u>CYCl</u> Promoters <u>With Both The</u>
<u>URA3 Gene And The LEU2 Gene on Expression Vectors</u></center>

      The construction of tripartite SMC expres-
sion vectors based on the <u>MFαl</u> secretion leader and
the <u>MFαl</u>, the <u>ACT</u>, and the <u>CYCl</u> promoters has been

25   described above. Each of these vectors contained
the yeast <u>URA3</u> gene for selection in yeast. We also
constructed vectors based on selection using the yeast
<u>LEU2</u> gene. The expression of the plasmid <u>LEU2</u> gene
is low due to a partial deletion; higher copy numbers

30   of the <u>LEU2</u> type expression vectors are needed to
allow the growth of transformants on leucine-free
media. E. Erhart and C. P. Hollenberg, "The Presence
of a Defective <u>LEU2</u> Gene on 2μ DNA Recombinant
Plasmids of <u>Saccharomyces cervisiae</u> is Responsible

35   for Curing and High Copy Number," <u>J. Bacteriol.</u>,
<u>156</u>, pp. 625-35 (1983). For a discussion of the use

of yeasts having the URA3 gene or the LEU2 gene in
recombinant DNA experiments, see D. Botstein et al.,
"Sterile Host Yeast (SHY): a Eukaryotic System for
Biological Containment for Recombinant DNA Experi-
5   ments," Gene, 8, pp. 17-24 (1979).

We constructed the LEU2-type vectors by
inserting the chosen promoter/MFα1 secretion leader/
SMC structural sequence into plasmid pJDB207, using
procedures similar to those used with the URA3
10   plasmids. The structures of the resulting plasmids
are shown in Figure 6. The plasmid pJDB207 is
described in J. D. Beggs, "Multiple-copy Yeast
Vectors," Molecular Genetics in Yeast, Alfred Benzon
Symposium 16, 383-95 (1981).

15   We transformed the three URA-type vectors
derived from plasmid pEX-2 (which is described in
J. F. Ernst and R. C. Chan, supra) and the three
LEU2-type vectors into yeast strain BJ1991. We
deposited two of the recombinant yeast strains in
20   the Deutsche Samulung Von Mikroorganismen (West German
Culture Collection). One strain, YE439, is a LEU2-
type vector with the ACT promoter/MFα1 secretion
leader/SMC structural sequence. It was deposited on
October 30, 1985 and is identified as DSM 3578.
25   Another strain, YE466, is a LEU2-type vector with
the CYC1 promoter/MFα1 secretion leader/SMC structural
sequence. It was deposited on October 30, 1985 and
is identified as DSM 3579. Transformants were grown
in the "SD" medium described in F. Sherman et al.,
30   "Methods in Yeast Genetics," Cold Spring Harbor
Laboratory, Cold Spring Harbor, New York (1981),
containing tryptophan and leucine (for URA3-type
vectors), or tryptophan and uracil (for LEU2-type
vectors) to an optical density of 2 at 600 nm. We
35   used these cultures to inoculate production medium,
consisting of SD medium containing 4% casamino acids
and tryptophan (inoculum was 10% of final volume of

production medium). At appropriate times during
growth, 0.5 ml of the cultures were pelleted using a
microfuge. We redissolved the cell pellets in 1/10
the original culture volume by boiling 5 min in sodium
5 dodecyl sulphate (SDS) sample buffer, which is
described in U. K. Laemmli, "Cleavage of the Struc-
tural Proteins During the Assembly of the Head of
Bacteriophage T4," Nature, 227, pp. 680-85 (1970).
We then determined the SMC levels in the redissolved
10 cell pellets and in the culture fluid after dilution,
using a radioimmune assay (Nichols Institute Diag-
nostics, San Juan Capistrano, California).
        With all six constructions, less than 10%
of the total SMC was cell associated. We purified
15 SMC present in the culture medium, and analyzed it,
finding that the amino terminal and carboxyl terminal
amino acids were identical to human SMC and that the
biological activity of the secreted SMC was the same
as human SMC.
20   Surprisingly, the weakest promoter used,
CYC1, resulted in the highest SMC secretion values
shown in Figures 3 and 7 and Table 2. This was
unexpected because the strongest promoter would have
been expected to give the highest secretion values.
25 Unexpectedly, the transformants carrying the CYC1
constructions grew slowest and showed a pronounced
optimum curve of SMC secretion, indicating the
occurrence of cell lysis and SMC degradation (see
Figure 7).
30   To determine the reason for the unexpected
results, we analyzed SMC plasmid copy number and
transcript levels in the transformants described
herein after three days in production medium. The
results are shown in Table 2. We discovered an
35 inverse relationship between copy number and promoter
strength in secretion systems. The secretion system
with the weakest promoter (CYC1) proved to have the highest

-27-

copy numbers. The MFα1 constructs had the lowest copy

numbers. The average SMC transcript level proved to be

proportional to copy numbers for both URA3 and LEU2

constructs. In turn, with LEU2 constructions, mRNA levels

5 were proportional to secretion levels. With URA3

constructions, no significant increase of SMC secretion

levels with increased mRNA was observed. While not wishing

to be bound by theory, the URA3 constructs do not appear

to have the same effect on the SMC secretion levels because

10 of a high percentage of URA3 plasmid loss.

A high percentage of URA3 transformants actually have

few or no URA3 expression plasmids. These unproductive

transformants contribute little to overall SMC secretion.

Conceivably, a high percentage of URA3 transformants also

15 have high copy numbers and high promoter strength. These

transformants would contribute little to SMC secretion

because of slow growth and reduced cell viability. This

suggested mechanism does not apply to LEU2 vectors because

the minimum copy number of the LEU2-type vectors in

20 transformants grown in leucine-free media is about 35

(Erhart and Hollenberg. supra). In addition, since the

production medium used contains leucine (but no uracil),

elevated copy numbers are not required for growth of LEU2

vector-transformants in this medium. Thus, yeast strains

25 transformed with LEU2 appear more homogeneous than URA3

vectors with respect to copy number and mRNA content.

With the URA3 transcripts, mRNA levels increased

-28-

as promoter strength weakened, even though secretion

levels remained fairly constant. This trend appears to

reflect the lower copy numbers for URA3 constructs. It

appears that, for LEU2 transcripts, a copy number of at

5   least 40 is required to obtain minumum levels of mRNA and

SMC secretion.

Table 2 does not give the actual ratio of SMC to

actin transcripts, and the numbers are for comparative

purposes only.

10   Transformation frequencies for all six plasmid

constructions are also shown in Table 2. Aproximately

similar transformation frequencies were obtained with all

the URA3-type vectors. However, dramatic differences

were seen for the LEU2-type vectors, the MFα1 promoter

15 constructions could not be transformed in yeast at all,

relatively high frequencies were observed for the ACT

promoter construction and lower frequencies of

transformation, compared to the control (pJDB207) and the

ACT construction, were characteristic of the CYC1 construction.

20 We did obtain one rare LEU2-type vector transformant with

the MFα1 construction; however, this transformant

contained a rearranged plasmid that directed the

synthesis of low amounts of SMC as shown parenthetically

in Table 2. This result suggests that the MFα1 promoter

25 does not permit a sufficiently high copy number to overcome

the effects of the LEU2 defect.

-29-

Thus, secretion rate, cell viability, promoter strength and vector copy number appear to be interrelated.

TABLE 2

Construction	plasmid	promoter	selection	transformants/µg	% loss	copy number	RNA	SMC (mg/l)
1	p336/1	MFα1	URA3	225	6	27	10.9	12
2	p364/1	ACT	URA3	319	10	32	26.3	12
3	p446/1	CYC1	URA3	296	12	49	40.2	13
--	p160/1	control	URA3	300	10	45	----	--
4	(pMF-SMC)	MFα1	LEU2	0(1)	(32)	(41)	(8.2)	(8)
5	p482/18	ACT	LEU2	4	0	69	31.9	20
6	p504/5	CYC1	LEU2	2	0	110	54.8	28
--	pJDB-207	control	LEU2	58	0	110	----	--

-31-

EXAMPLE 5

Secretion of TNF

We constructed fusions of the MFα1 secretion
leader to the gene encoding human tumor necrosis
5   factor (TNF) as shown in Figure 8 by using the pro-
cedure described above for MFα1/SMC fusions. We then
analyzed the effect of promoter replacement on TNF
secretion.

We ligated the following three fragments
10   to construct vector pTNF11 (see Figure 8): (1) a
1.2 kb EcoRI-StuI fragment carrying the promoter and
the secretion leader of MFα1 ending in a StuI site
(described supra); (2) a 9 kb EcoRI-HindIII fragment
of pEX-2 carrying the yeast 2μ origin of replication
15   and the URA3 gene for selection in yeast (described
supra); and (3) a blunt-ended HindIII fragment
(0.9 kb) carrying the TNF gene. Isolation procedures
for this third fragment are disclosed in D. Pennica
et al., "Human Tumor Necrosis Factor: Precursor
20   Structure, Expression and Homology to Lymphotoxin,"
Nature, 312, pp. 724-29 (1984). We generated the
blunt-ended HindIII fragment by using synthetic
oligonucleotide linkers to recreate the 5' end of
the TNF gene by extension from an AvaI site situated
25   close to the 5' end of the TNF gene (valine is the
first amino acid of TNF). The junction of MFα1 to
the TNF gene had the following sequence:

               (5')...AAA AGG GTA CGT TCT TCC...(3')
                      lys arg val arg ser ser
30   MFα1   TNF

The construction of a TNF secretion vector
based on the ACT promoter, instead of the MFα1 pro-
moter, is also shown in Figure 8. We ligated the
following fragments to construct pACT-TNF-EX: (1) a
35   1.2 kb PstI-HindIII fragment from pTNF11 carrying
most of the MFα1 secretion leader fused to the TNF

gene (described supra); (2) a 0.5 kb BamHI-PstI frag-
ment from p355/1, shown in Figure 5, carrying the.
ACT promoter and part of the MFα1 leader (described
supra); and (3) the 9 kb BamHI-HindIII fragment of

5   pEX-2 (described supra).

We transformed yeast strain BJ1991 with
pTNF11 or pACT-TNF-EX and selected for Ura^{+} proto-
trophs. Transformants were grown selectively in
liquid SD medium lacking uracil. Shake flasks con-

10   taining YPD medium, which is disclosed in F. Sherman
et al., supra, were inoculated with minimal cultures
(10% final volume) and incubated at 30°C on a rotatory
shaker. At appropriate times during incubation, we
centrifuged 1 ml samples of the yeast culture for . .

15   1 min using a microfuge. We then treated the cell
pellet with zymolase to generate spheroplasts. We
then lysed the spheroplasts using 1% triton X-100 at
20% of the original culture volume, followed by
removal of cell debris by a short centrifugation step.

20   ·   We electrophoresed samples containing TNF
on 12.5% SDS-acrylamide gels and transferred the
proteins to nitrocellulose by standard methods. We
probed the protein blot with rabbit antibodies for
human TNF, and we visualized regions of the blot

25   containing TNF/anti-TNF complexes using peroxidase-
coupled swine anti-rabbit antibodies. Maximal TNF
expression results in the culture fluid and in cell
extracts are shown in Table 3. The results indicated
that a significant improvement of secretion of TNF

30   was obtained when the MFα1 promoter was replaced by
the ACT promoter. Only 10-20% of the total TNF pro-
duced was cell associated.

-33-

TABLE 3

MAXIMAL TNF EXPRESSION LEVELS BY YEAST TRANSFORMANTS

plasmid	cells (mg/litre)	medium (mg/litre)
pTNF11	1	5
pACT-TNF-Ex	1.5	15

EXAMPLE 6

Secretion of TPA

The gene coding for human tissue plasminogen
activator (TPA) carries a convenient BglII site at a
position corresponding to the first amino acid of
TPA. D. Pennica et al., "Cloning and Expression of
Human Tissue-type Plasminogen Activator cDNA in
E. coli," Nature, 301, pp. 214-21 (1983). In order
to allow correct fusion to the MFα1 secretion leader,
we introduced a convenient BglII site at the MFα1
position corresponding to the lys-arg junction. The
structure of the MFα1/TPA fusion point in the
following constructions is as follows:

```
(5')...AAA AGA TCT TAC...(3')
       lys arg ser tyr
       MFα1      TPA
```

We inserted a BglII fragment carrying the
MFα1 promoter and secretion leader into the single
BglII site of the previously constructed vector pPAY4
as shown in Figure 9. In this manner, the PHO5 pro-
moter and secretion leader were replaced by the MFα1
promoter and secretion leader. We further modified
the resulting expression vector, pMF-TPA, by replacing
the PHO5 and MFα1 promoters by the ACT promoter, as
shown in Figure 9, to obtain expression vector pACT-
MF-TPA. The PHO5 promoter present in both TPA
constructions is repressed in regular (high-phosphate)
production media. Accordingly, the expression results
described below are due to the activity of the MFα1

-34--

promoter because the tests were carried out in regular
(high phosphate) media.

We used each of the expression plasmids,
pACT-MF-TPA and pMF-TPA, to transform yeast cells of
5   strain BJ1991, as described above, selecting for
Leu[+] prototrophs. We selectively grew transformants
in SD medium. We inoculated shake flasks containing
YPD medium with the SD cultures (10% of final volume)
and incubated at 30°C. At appropriate times during
10   incubation, we prepared cell extracts as described
above in Example 4.

We determined TPA activity in cell frac-
tions by halo formation on fibrinogen-plasminogen-agar
plates as described in A. Granelli-Piperno and
15   E. Reich, "A Study of Proteases and Protease-Inhib-
itor Complexes in Biological Fluids", J. Exp. Med.,
148, pp. 223-34 (1978). The results shown in Table 4
indicate that replacement of the MFα1 promoter by
the ACT promoter is advantageous in obtaining secre-
20   tion of heterologous proteins by yeast using alpha
mating factor fusions.

TABLE 4

MAXIMAL TPA EXPRESSION LEVELS BY YEAST TRANSFORMANTS

plasmid	cells (μg/litre)	medium (μg/litre)
pMF-TPA	30	0
pACT-MF-TPA	100	5

As shown in Table 3, most of the TPA pro-
duced was cell associated; only 5% of the total TPA
30   produced appeared in the medium. However, the
following evidence suggested that yeast TPA had
entered the secretion pathway: ·

(1)   the TPA expressed was biologically
active, indicating correct folding. Attempts

to express TPA in yeast without a secretion signal sequence produced non-active TPA; and

     (2)   the cell associated TPA was glyco-sylated. Secretion at least into the lumen of
5   the endoplasmic reitculum is necessary for glycosylation to occur.

     It will be apparent to those skilled in the art that various modifications may be made in the invention without departing from its spirit or
10   scope, and our basic construction can be altered to provide other embodiments which utilize the processes and compositions of this invention. Therefore, it will be appreciated that the scope of this invention is to be defined by the claims appended hereto rather
15   than the specific embodiments which have been presented as examples.

-36-

WHAT IS CLAIMED IS:

1.   A DNA sequence comprising:

(a)   a promoter of at most intermediate strength; and

(b)   a DNA secretion signal sequence, beginning with a start codon, heterologous with respect to said promoter and operatively linked to said promoter.

2.   The DNA sequence of claim 1, wherein said start codon is ATG.

3.   The DNA sequence of claim 1 or 2, wherein said promoter is selected from the group consisting of: the ACT, CYC1, CYC7, trp1, URA3, LEU2, trp and lac promoters.

4.   The DNA sequence of any one of the preceding claims, wherein said DNA secretion signal sequence is the MFα1 secretion signal sequence.

5.   A DNA sequence according to any one of the preceding claims further comprising a DNA sequence coding for a desired protein operatively linked to said DNA secretion signal sequence.

6.   The DNA sequence of claim 5, wherein said desired protein is capable of being secreted by a microorganism or cell.

7.   The DNA sequence of claim 5 or 6, wherein said desired protein is selected from the group consisting of serum proteins, analgesic poly-peptides, β-endorphin, somatostatin, SMC, insulin, human growth hormone, bovine growth hormone, luteinizing hormone, ACTH, pancreatic polypeptide

preproteins, TPA, TNF, preproinsulin, proinsulin, the A chain of insulin and the B chain of insulin.

8. A multicopy expression vector comprising a DNA sequence according to any one of
5   claims 1-7.

9. The multicopy expression vector of claim 8, wherein said multicopy expression vector is selected from the group consisting of plasmids and multicopy phages.

10   10. A host transformed by the DNA sequence of any one of claims 1-7.

11. A host transformed by the multicopy expression vector of any one of claims 8-9.

          12. The host of any one of claims 10-11,
15   wherein said multicopy expression vector is selected from the group consisting of: ARS-type plasmids, 2μ-type plasmids, ColEl type plasmids, pBR322, RP4 plasmid, M13 phage and lambda phage.

          13. The host of any one of claims 10-12,
20   wherein said promoter does not materially reduce the copy number of said multicopy expression vector in said host.

          14. The host of any one of claims 10-13, wherein the expression control sequence of said
25   promoter produces less than 0.3 percent of the total mRNA within said host.

15.   The host of claim 14, wherein at most 0.15 percent of the total mRNA within said host is the product of said promoter.

16.   The host of any of claims 10-15, wherein said host is selected from the group consisting of: bacteria, fungi, yeasts, animal cells, plant cells in tissue culture or human tissue cells.

17.   The host of any one of claims 10-16, wherein said host is selected from the group consisting of: E.coli, Bacillus, Streptomyces, Pseudomonas, S. cerevisiae, CHO cells, mouse cells, swine cells, bovine cells, fowl cells or fish cells.

18.   The host of any one of claims 10-17, wherein said host is yeast.

19.   The host of claim 18, wherein said yeast is S. cerevisiae.

20.   A method for producing a desired protein, comprising culturing the host of any one of claims 10-19.

Figure 1

EcoRI

A

pYA301
8 kb

*Xho*I

*Bam*HI
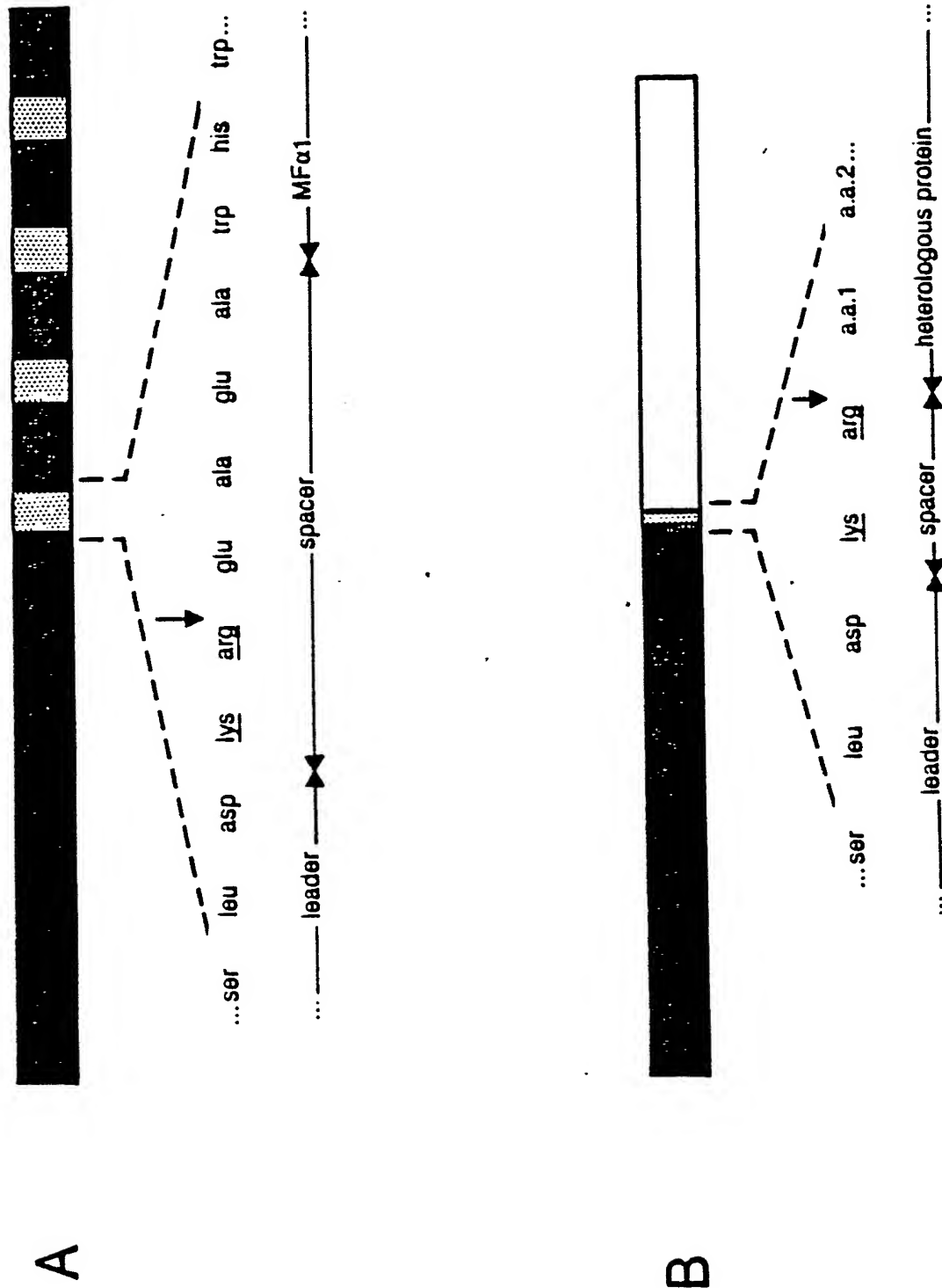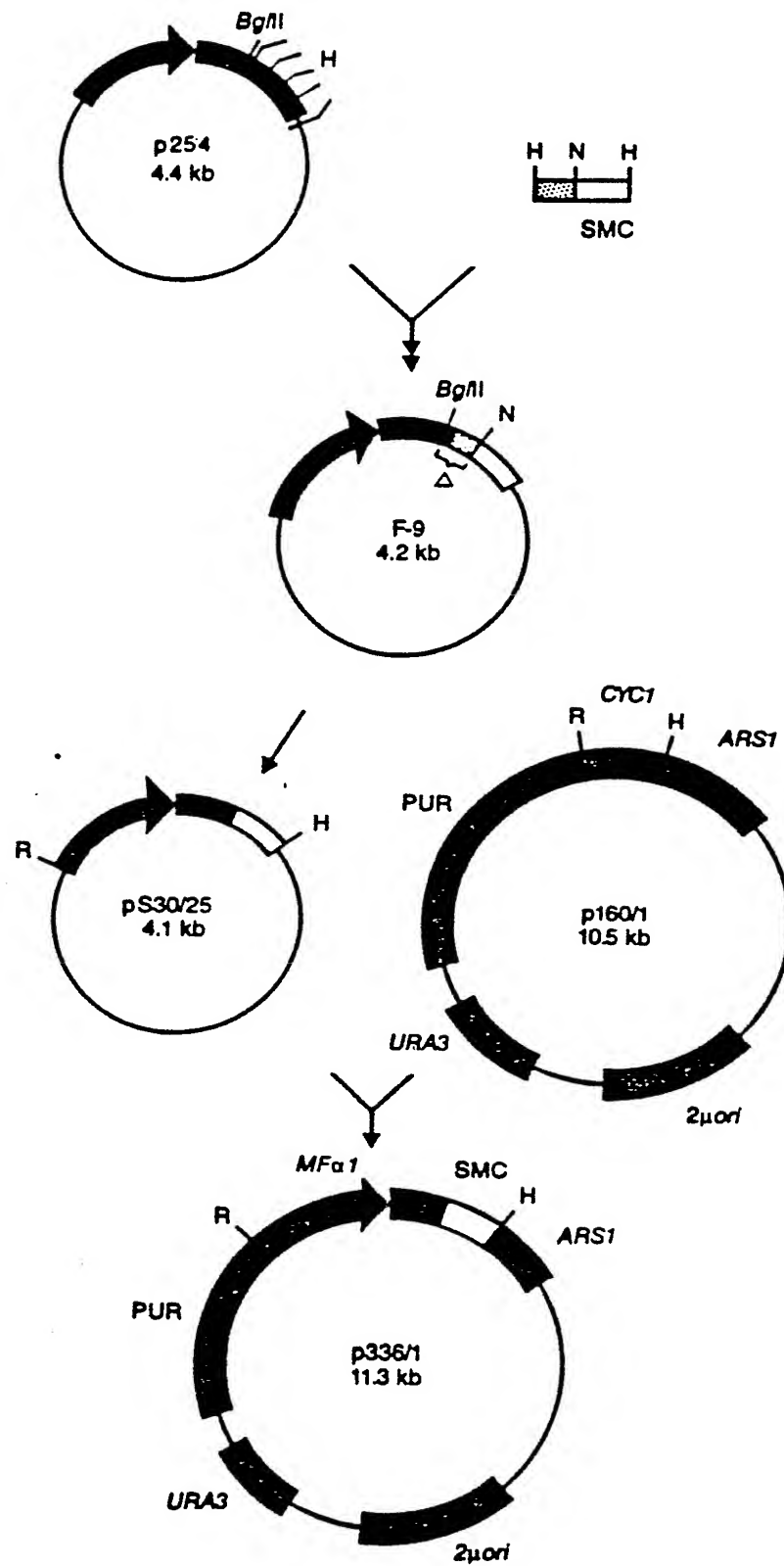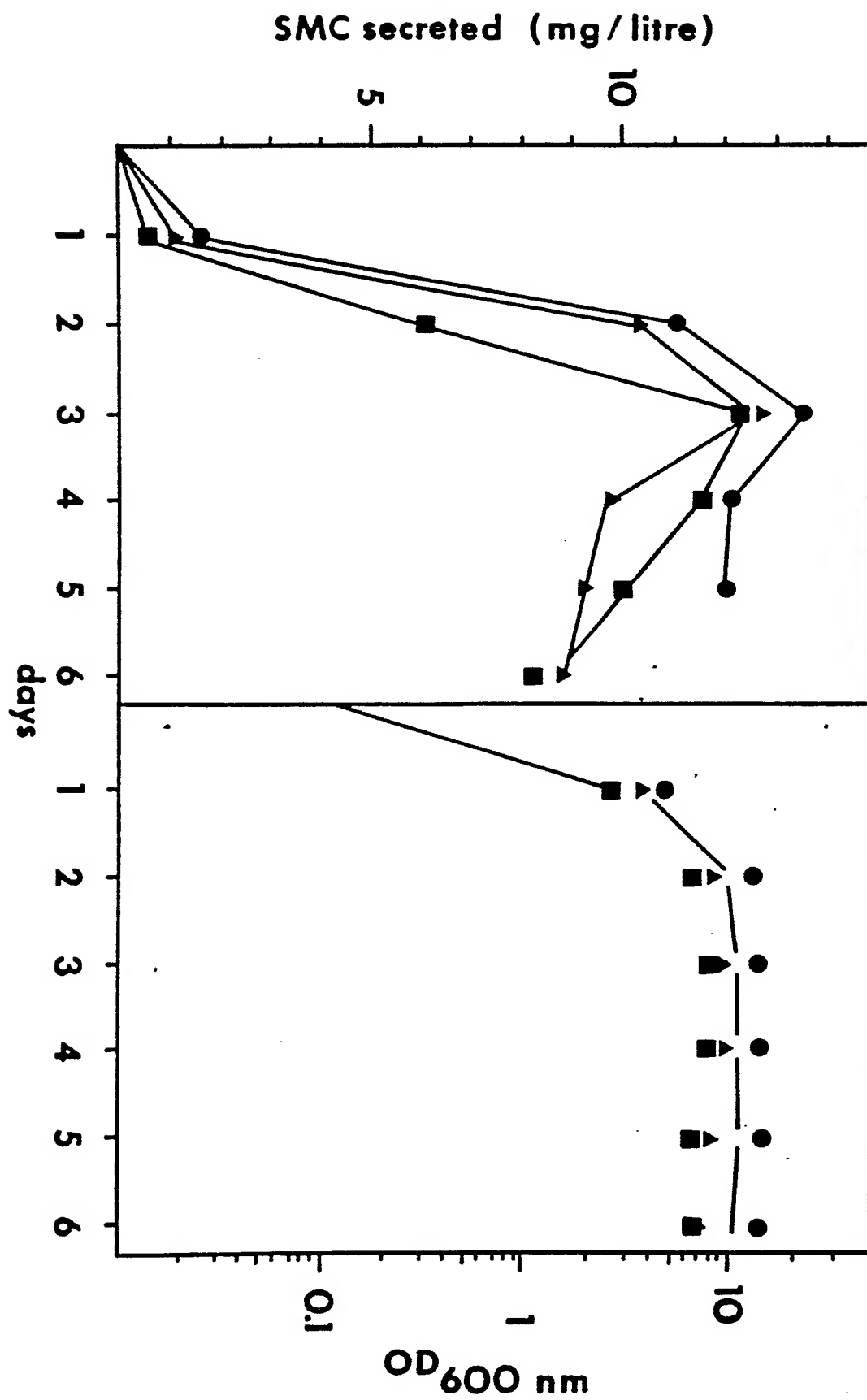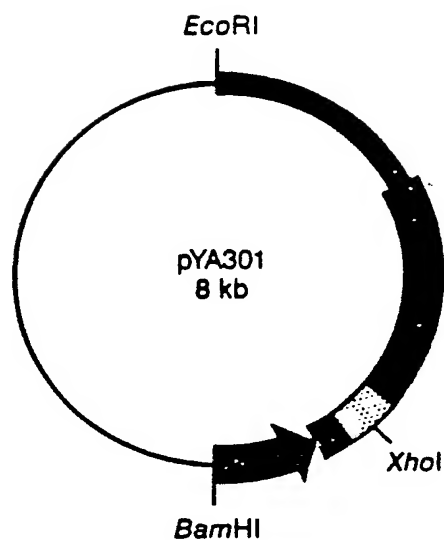
1. digest with *Xho*I
2. treat with Bal31
3. digest with *Eco*RI
4. fill-in with dNTPs
5. religate

B

pΔ4
~ 4.5 kb

*Bam*HI   *Eco*RI
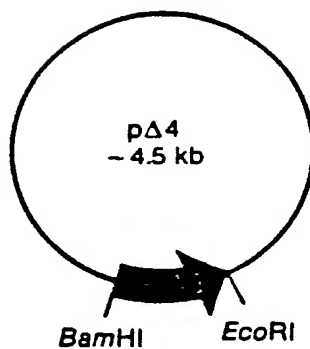
```
          -20             -10             +1
Actin  GATCGAAAATTTACTGAATTAACAATGGATTCTGG
pEX-5  GAATTC
pEX-7  GATCGAAAATTTACTGAATTC
pEX-8  GATCGAAAATTTACTGAATTAACAATGGATTCTGAATTC
```

Figure 5


isolate large *PstI - HindIII* fragment

Figure 7

Figure 8

MFα1
BglII   BglII

B   PHO5   BglII
                    TPA
                    H

2μori

pPAY4
10.5 kb

LEU2

BglII   MFα1   PstI
                    BglII
B   TPA
                    H

2μori

pMF-TPA
11.7 kb

LEU2

1. isolate large BamHI - BglII fragment
2. isolate 0.23 kb BglII - PstI fragment

p355/1

B   P

ACT
                    TPA
                    H

2μori

pACT-MF-TPA
10.4 kb

LEU2
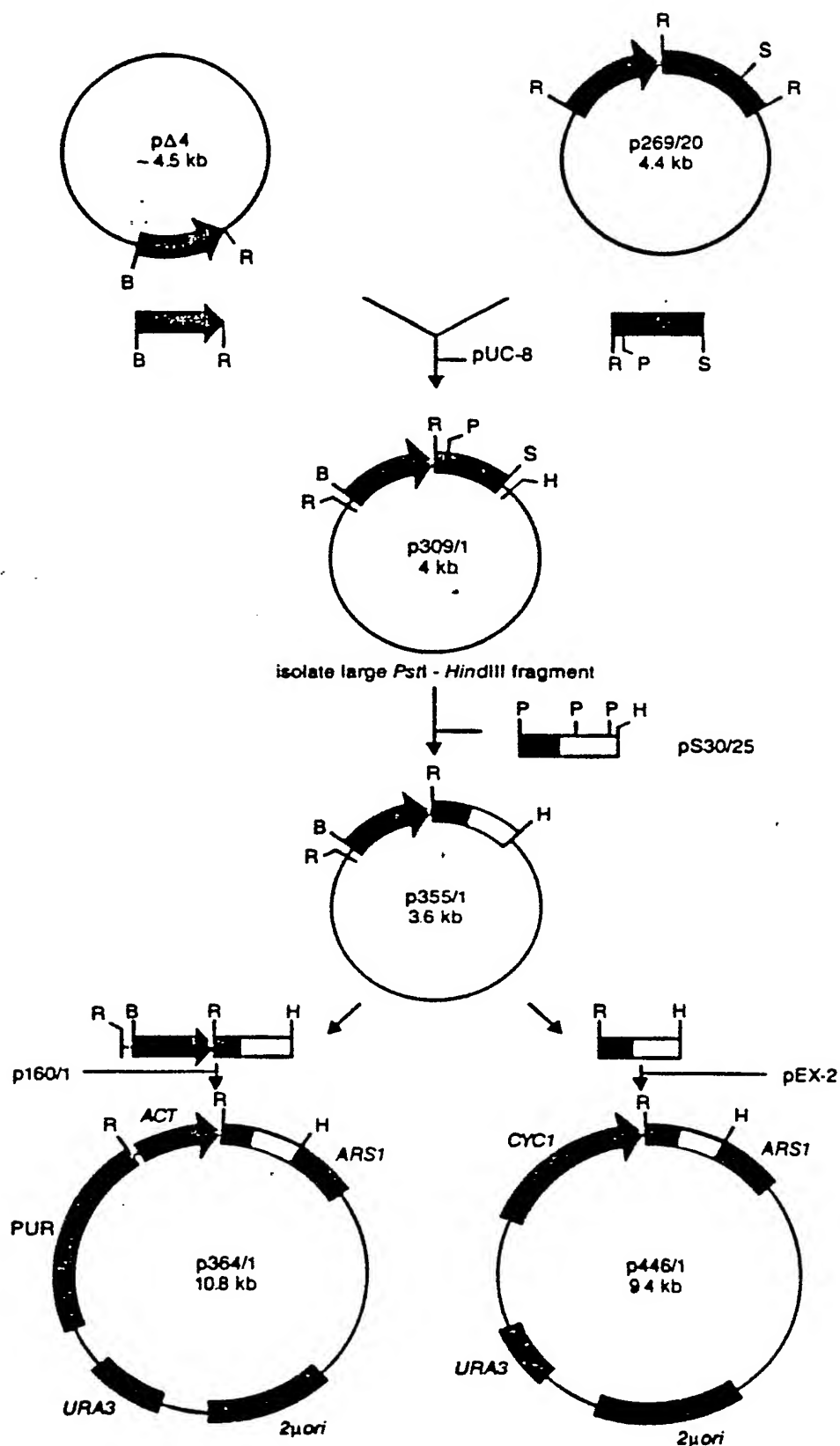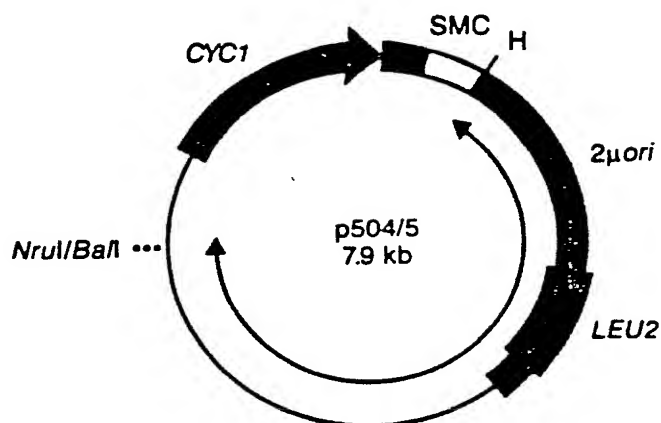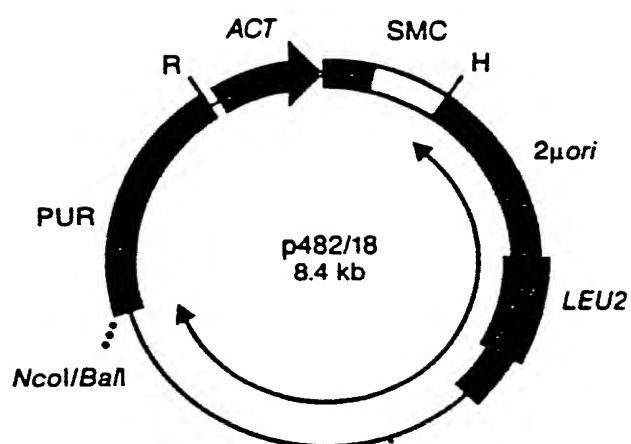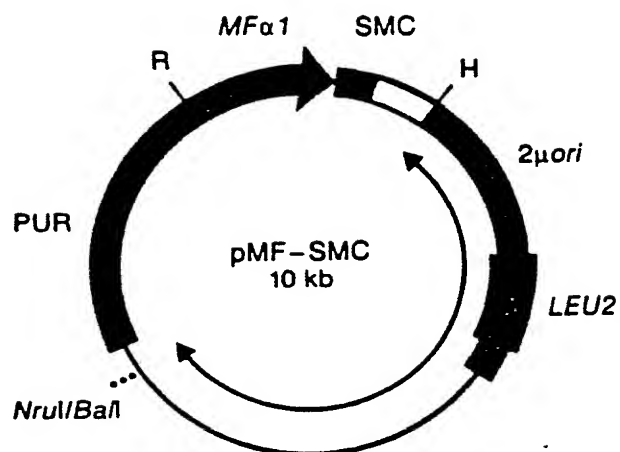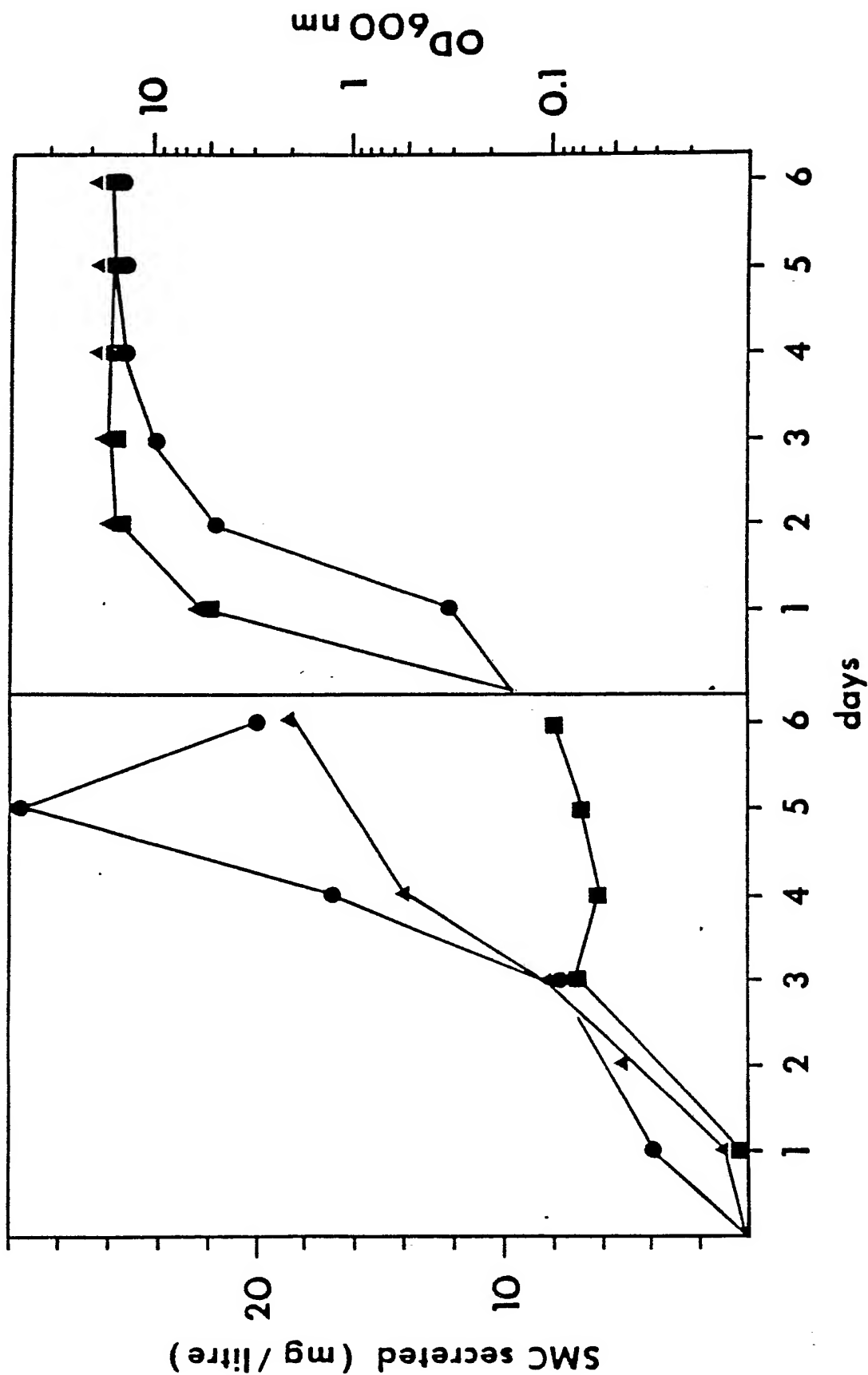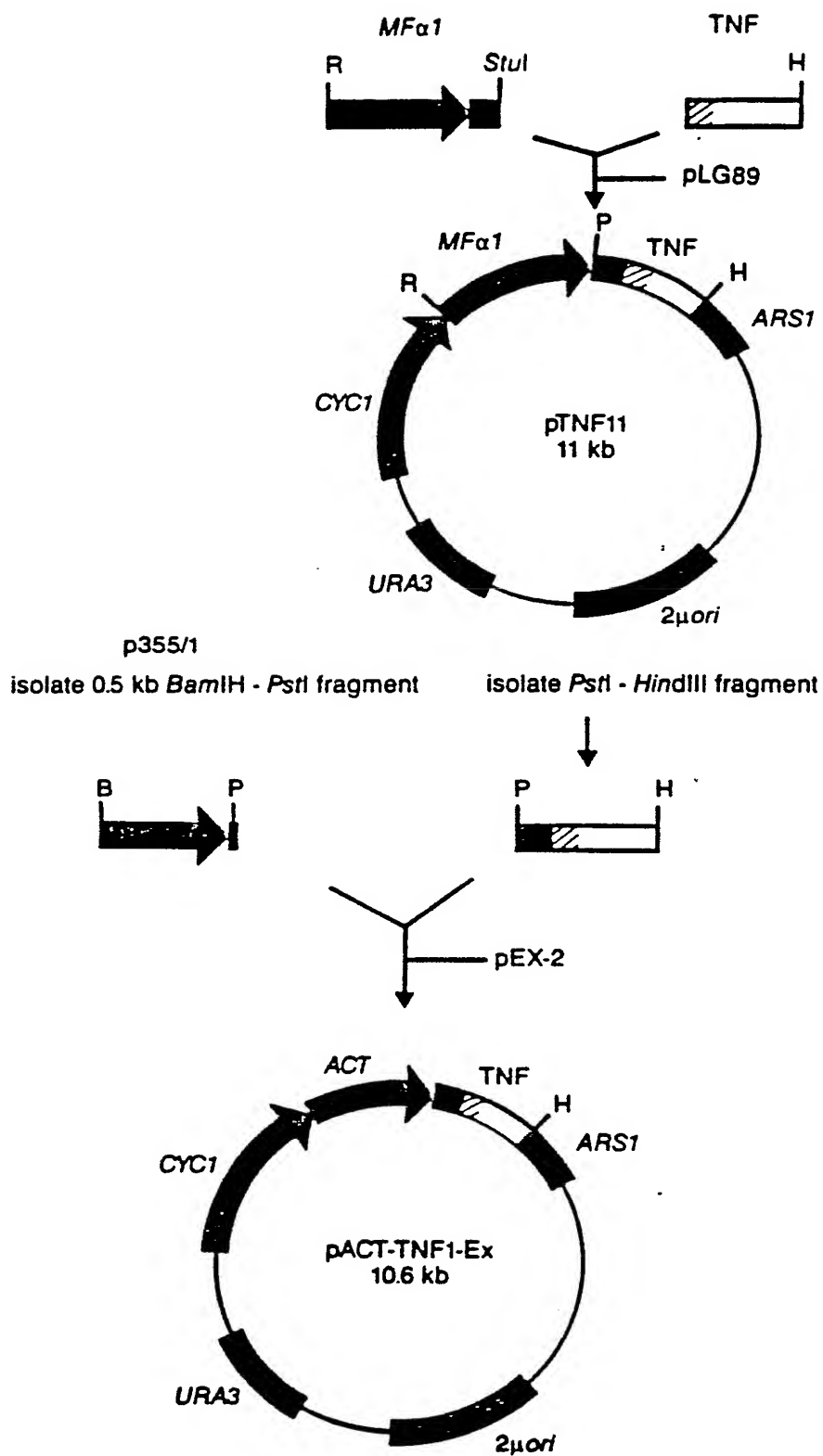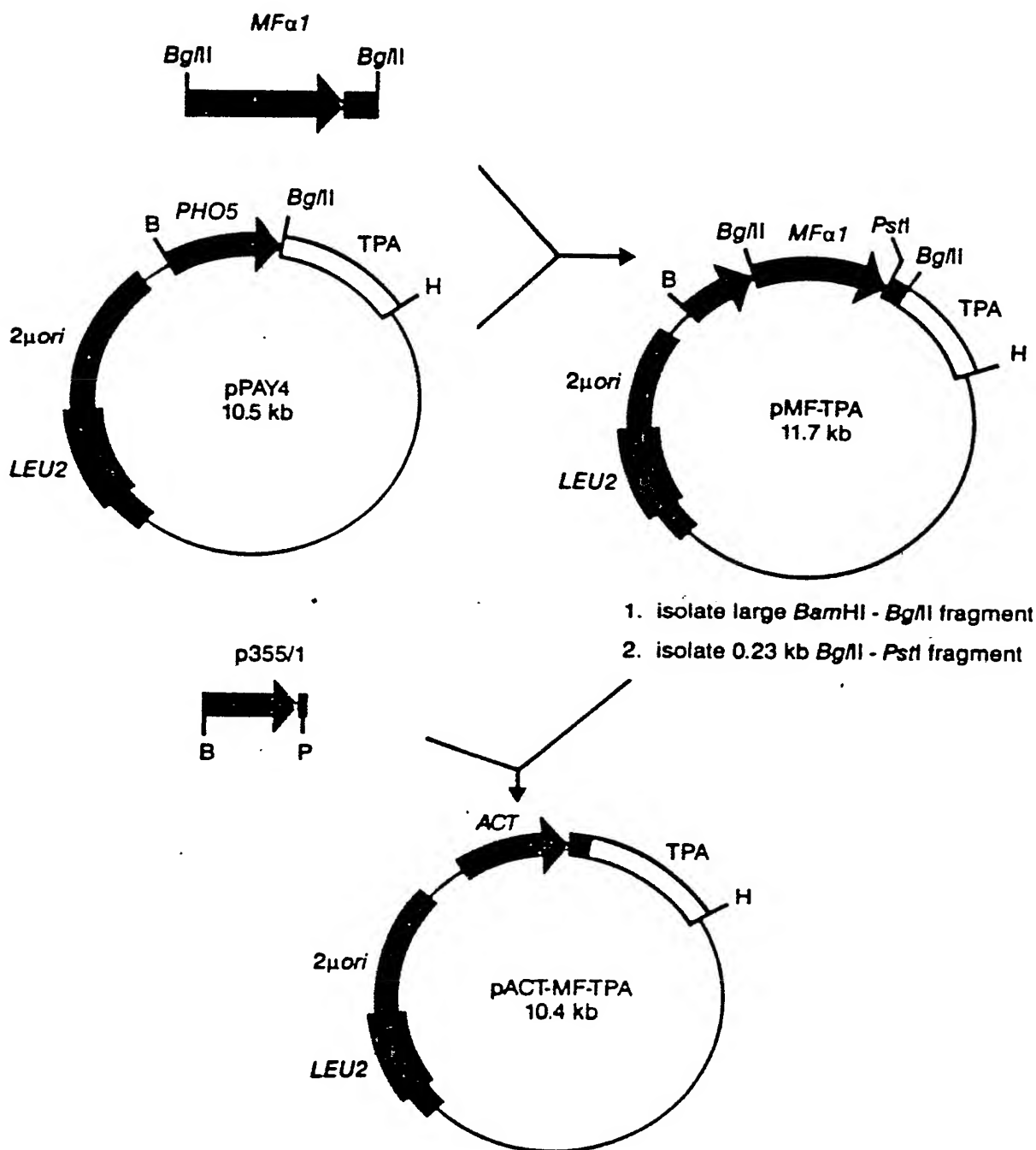
INTERNATIONAL SEARCH REPORT

International Application No PCT/EP 86/00675

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) *

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC^4: C 12 N 15/00; C 12 N 1/18; C 12 N 1/20; C 12 N 5/00
       C 12 P 21/00

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC^4	C 12 N; C 12 P

Documentation Searched other than Minimum Documentation
to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category *	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
X	Nucleic Acids Research, volume 11, no. 8, 1983, M.J. Dobson et al.: "Expression in saccharomyces cerevisiae of human interferon-alpha directed by the TRP1 5' region", pages 2287-2302 see page 2295, paragraph 2; page 2296 --	1-3,5,8-20
X	Gene, volume 35, 1985, M.A. Peters et al.: "Expression of a biologically active analogue of somatomedin-C/insulin-like growth factor I", pages 83-89 see the whole document --	1,3,5,7-13, 16,17,20
X	Nature, volume 287, 2 October 1980, D.V. Goeddel et al.: "Human leukocyte interferon produced by E. coli is biologically active", pages 411-416 see page 414, column 2 --	1-3,5,8-13, 16,17,20
X,P	EP, A, 0200590 (GENETICA) 5 November 1986 see page 16, lines 22-29	1-3,5,8-13, 16,17,20

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
25th February 1987	-2 APR 1987
International Searching Authority	Signature of Authorized Officer
EUROPEAN PATENT OFFICE	M. VAN ...

Form PCT/ISA/210 (second sheet) (January 1985)

This Annex lists the patent family members relating to the
patent documents cited in the above-mentioned international
search report. The members are as contained in the European
Patent Office EDP file on 10/03/87

The European Patent Office is in no way liable for these
particulars which are merely given for the purpose of
information.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A- 0200590	05/11/86	FR-A- 2579224	26/09/86
		JP-A- 61275223	05/12/86